**Blood First Edition Paper, prepublished online January 9, 2003; DOI 10.1182/blood-2002-06-1796**

# Alternative 5' exons and differential splicing regulate expression of protein 4.1R isoforms with distinct N-termini

[1]Marilyn K. Parra, [1]Sherry L. Gee, [2]Mark J. Koury, [3]Mohandas Narla, [1]John G. Conboy

[1]Lawrence Berkeley National Laboratory, Life Sciences Division, Berkeley, CA 94720; [2]Dept. of Medicine, Vanderbilt University, Veterans Affairs Medical Centers, Nashville, TN 37232; [3]New York Blood Center, New York, NY 10021

Running head: Regulation of protein 4.1R expression

Scientific heading: Red Cells

Research support: This work was supported by NIH grants HL45182 and DK32094, by the Director, Office of Biological and Environmental Research, US Department of Energy under contract DE-AC03-76SF00098, and by a Merit Review Award from the Department of Veterans Affairs to MJK.

Word counts: Abstract, 250 words
Total text, 4829 words (excludes title page, references, figure legends)

Corresponding author: John G. Conboy
Lawrence Berkeley National Laboratory
Life Sciences Division
Mailstop 74-157
1 Cyclotron Road
Berkeley, CA 94720
(510) 486-6973
(510) 486-6746 (fax)
jgconboy@lbl.gov

## Abstract

Among the alternative pre-mRNA splicing events that characterize protein 4.1R gene expression, one involving exon 2' plays a critical role in regulating translation initiation and N-terminal protein structure. Exon 2' encompasses translation initiation site AUG1 and is located between alternative splice acceptor sites at the 5' end of exon 2; its inclusion or exclusion from mature 4.1R mRNA regulates expression of longer or shorter isoforms of 4.1R protein, respectively. The current study reports unexpected complexity in the 5' region of the 4.1R gene that directly affects alternative splicing of exon 2'. Three mutually exclusive alternative 5'exons, designated 1A, 1B, and 1C, were identified far upstream of exon 2 in both mouse and human genomes; all three are associated with strong transcriptional promoters in the flanking genomic sequence. Importantly, exons 1A and 1B splice differentially with respect to exon 2', generating transcripts with different 5' ends and distinct N-terminal protein coding capacity. Exon 1A-type transcripts splice so as to exclude exon 2' and therefore utilize the downstream AUG2 for translation of 80kD 4.1R protein, whereas exon 1B transcripts include exon 2' and initiate at AUG1 to synthesize 135kD isoforms. RNA blot analyses revealed that 1A transcripts increase in abundance in late erythroblasts, consistent with the previously demonstrated upregulation of 80kD 4.1R during terminal erythroid differentiation. Together these results suggest that synthesis of structurally distinct 4.1R protein isoforms in various cell types is regulated by a novel mechanism requiring coordination between upstream transcription initiation events and downstream alternative splicing events. (jgconboy@lbl.gov)

## Introduction

A family of protein 4.1R polypeptides is encoded by a single complex locus on human chromosome 1. Best characterized among the 4.1R isoforms is an 80kD polypeptide that functions as a key structural component of the membrane skeleton in mature red cells. Protein 4.1R's critical role in maintaining the specialized mechanical properties of the red cell plasma membrane is demonstrated by the abnormal morphology and increased membrane fragmentation of 4.1R deficient red cells, occurring as a result of natural mutations in humans [1,2] or by gene knockout experiments in mice [3]. A considerable body of work has defined a network of protein-protein interactions by which 4.1R participates in

stabilization of the spectrin-based membrane skeleton, and in connection of this skeleton to overlying plasma membrane. Moreover, recent crystal structure data has provided new insights into the structure and interactions of the membrane binding domain [4].

In contrast to the relatively simple expression pattern of 4.1R in mature red cells, a more complex array of 4.1R protein isoforms of varying sizes and proposed functions have been reported in erythroid progenitors and nonerythroid cells. Early erythroid progenitors, and many nonerythroid cells, express 4.1R isoforms differing in size and subcellular localization from the familiar erythroid form. Among the sites of function proposed for nonerythroid 4.1R are the nucleus, centrosomes, and spindle poles of dividing cells [5-9]; epithelial tight junctions [10]; contractile sarcomeres in muscle [11]. Characterization of 4.1R knockouts in mice indicates a role in selected neurons that affect fine motor control [12]. Finally, 4.1R interactions have been reported with a variety of other proteins, including translation and splicing factors [13-21], implying that 4.1R may have additional roles in the cell. The concept that 4.1R has a diverse set of cellular functions in addition to its role in red cells is thus well established. However, much remains to be learned about the structure and function of these 4.1R isoforms in nucleated cells.

A major challenge in the field is to understand how a single 4.1R gene can express such a variety of isoforms with diverse structures and functions. Molecular characterization of this complex gene has shown that it is ~240kb in length ([22,23] and our unpublished results) and is subject to extensive regulation at the level of alternative pre-mRNA splicing (reviewed in [24]). At least ten of the internal coding exons of the gene can be alternatively spliced, and several of these are tightly regulated in tissue- or developmental-specific patterns. As a direct consequence, different cell types can express different profiles of 4.1R mRNA isoforms [11,25], and thereby synthesize functionally different complements of 4.1R protein. Of particular relevance to erythroid differentiation and red cell function is alternative exon 16, which encodes an essential part of the spectrin-actin binding domain. Exon 16 is excluded in early erythroid progenitors but included in later progenitors, allowing the mature cells to produce 4.1R isoforms that can effectively stabilize the developing membrane skeleton [26,27].

The current study focuses on novel aspects of protein 4.1R gene structure that play a critical role in regulating the balance in synthesis of 80kD vs 135kD 4.1R protein isoforms. Specifically, we report an unexpected complexity in 5' exon/intron organization: three mutually exclusive 5' exons map far upstream of the coding exons, and each of these alternative exons possesses its own transcriptional

promoter, suggesting that they may represent alternative first exons of the 4.1R gene. Most importantly, exons 1A and 1B splice differentially to two acceptor sites downstream in exon 2' in a manner that regulates expression of translation initiation site AUG1. Transcripts initiated at exon 1A splice to an internal acceptor site in exon 2, thereby skipping exon 2' sequences and translation initiation site AUG1; the resulting spliced mRNAs encode 80kD isoforms of 4.1R protein by initiation at downstream AUG2 in exon 4. Conversely, transcripts initiated at exon 1B splice to include exon 2', incorporating AUG1 into the mature mRNA and facilitating synthesis of 135kD protein 4.1R. These results suggest that there is a mechanism for functional coupling between upstream transcription and downstream alternative splicing events.

**Materials and methods**

Genetic database analyses. Identification of alternative 5' exons in the 4.1R gene required analysis of both cDNA and high throughput genomic sequence (htgs) information in Genbank. To identify candidate first exons, we used the human 4.1R exon 2 sequence as a probe for BLAST searches of Genbank. Among several human and mouse cDNA clones that contained sequence information upstream of exon 2, three unique sequences were found to splice properly to exon 2. These candidate exons were designated as exon 1A, exon 1B, and exon 1C. Exon 1A was identical to the 5' end of the original human 4.1R cDNA (M61733), exon 1B was identical to mouse 4.1R cDNA L00919, and exon 1C represented unique sequence. In order to confirm that these were bona fide exons in the 4.1R gene, these sequences were used as probes in a BLAST search of the htgs database. All three sequences were found in BAC clones upstream of known coding exons of the 4.1R gene. Of note, a human ortholog of mouse exon 1B was characterized by 79% identity over a region of 107nt (see Figure 1A). Although exon 1B has not been found in any human cDNA clones in the databases, its authenticity was confirmed by RT-PCR experiments demonstrating in many tissues the presence of 4.1R transcripts possessing exon 1B sequences properly spliced to exon 2 (see Figure 3). In the current human genome assembly (June 2002), the coordinates for exon 1A are 29,219,768-29,219,837; for exon 1B: 29,226,261-29,226,464; for 1C: 29,247,204-29,247,395; and for exon 2: 29,320,058-29,320,532.

Amplification of 1A and 1B transcripts from human tissues. Splicing patterns of 1A- and 1B type transcripts were characterized using RT-PCR techniques to amplify 4.1R mRNA from several different human tissue sources (Clontech). One μg of total RNA was transcribed into cDNA using a specific antisense primer in exon 2 in a total volume of 10μl. Then 2μl of cDNA was amplified as described

previously [28] using the following primers: 1A-S, 5'-GCAAAGTGGCAGGAACCTCTTAAAG-3' and 2-AS: 5'-CGAGGAGAATAGTCGTGAAAGTCC-3', or 1B-S: 5'-GACTGGCTGCGTGACCCCGACGGCTG-3'; and 2-AS. Thirty-five cycles of amplification were performed with the use of an automated Perkin-Elmer Cetus 2400 thermal cycler under the following conditions: denaturation for 20 seconds at 94°C; annealing for 20 seconds at 60°C; extension for 40 seconds at 72°C. DNA fragments were analyzed by 5% polyacrylamide gel electrophoresis. The identity of all major PCR products discussed in this paper was confirmed by DNA sequence analysis.

Construction of PCR standards. Two potential 5' cDNA structures, containing exon 1A spliced to the proximal 3' splice site in exon 2 ("1A-proximal"), or exon 1B spliced to the distal 3' splice site in exon 2 ("1B-distal"), were never amplified from natural mRNA extracted from human tissues. However, it was essential to generate PCR size standards for these products in order to verify that such products would be detectable if they were present in a tissue. For the 1A-proximal product, we found that in vitro splicing of an artificial pre-mRNA construct (containing only exon 1A, a dramatically truncated intron, and exon 2) generated both 1A-proximal and 1A-distal products. A different approach employing splice overlap techniques [29] was required to generate the 1B-distal marker. Exon 1B was amplified using primer 1B-S (see above) and primer 2/1B-AS: 5'-CGGCCTCAGTCACTAAACTCTTCT-CCAGAAGCAGCCGTCGGGGTCAC-3'. Exon 2 (excluding 2') was amplified using primer 1B/2-S: 5'-GTGACCCCGACGGCTGCTTCTGGAGAAGAGTTTAGTGACTGAGGCCG-3' and primer 2-end: 5'-GGCC*TCTAGA*CTGTGTTTCTGCACTGCTTAAT-3' (italicized region is an XbaI restriction site for cloning purposes). Underlined portions of the primers created a 24bp overlap between the fragments, which were mixed in equal proportions and amplified with the outside primers 1B-S and 2-AS to create the 1B-distal marker.

RNA blots. Northern blot analysis of RNA was performed as previously described, using 10μg of total RNA per lane. The full length mouse 4.1R probe was prepared by amplification with the following primers: forward, 5'-ATGACAACAGAGAAGAGTTTAGTGGCTGAAGC-3'; reverse, 5'-TCACTC-CTCAGAGATCTCTGTCTCCTGGTGGACGACC-3'. The mouse exon 1A probe was also prepared by amplification with primers: forward, 5'-TCAGTGCAGGTGGAGGCCCCCGCGGG-3'; reverse, 5'-CTGGGCTGCACAGGACAGGGACCTGG-3'. The expression patterns of exons 1A, 1B, and 1C among human tissues was determined by hybridization to master RNA blots containing poly $A^+$ RNA from a number of different tissues (Clontech, Palo Alto, CA). Probes used for RNA dot blots were

prepared by PCR amplification of exon specific DNA fragments using the following primers: exon 1A, forward 5'-AAAGGGCGAGAGCGGCGCGGAG-3', reverse 5'-GCTGGGTGCGGCGGGGA-3'; exon 1B, forward 5'-TGCTGTCATTTGCTTCCATGTG-3', reverse 5'-CGTCGGGGTCACGCAGCCAGTC-3'; exon 1C forward 5'-CTCCCTCCGCTGGTGCGTTTAG-3', reverse 5'-ACGAACAATGGCAGGAGGAG-3'.

Erythroblast cell procurement and culture. Erythroid cells obtained from the spleens of mice infected with the anemia-inducing strain of Friend erythroleukemia virus were isolated and cultured as previously described [30,31]. Cells at t=0 hour are mainly proerythroblasts, which then differentiate over ~48hrs into late-stage erythroblasts and enucleated reticulocytes. Total RNA was isolated from cell pellets using RNeasy columns according to the manufacturer's instructions (Qiagen,Valencia, CA).

Transcriptional promoter assays. A 1.5kb region of genomic DNA upstream of exon 1A, and overlapping by ~50nt the exon 1A sequence shown in Figure 1A, was cloned upstream of the firefly luciferase reporter in pGL2-basic. Analogous regions upstream of exons 1B and 1C were similarly cloned into pGL2-basic. For each construct, transcriptional activity of the putative promoter region was assessed following transfection into HEK293T (human embryonic kidney) cells by measurement of luciferase activity in comparison to the activity of a control promoterless construct, pGL2basic. As a positive control we employed the pGL2-promoter vector containing an SV40 promoter (Promega, Madison, WI). Cells were co-transfected with a Renilla luciferase pRL-TK reporter to provide an internal control value to which the experimental firefly luciferase measurements can be normalized, to account for potential differences in transfection efficiency. Firefly and Renilla luciferase activities were measured using a dual luciferase assay system (Promega, Madison, WI) according to the manufacturer's instructions.

**Results**

Evidence for multiple alternative 5' exons in the Protein 4.1R gene. The existence of multiple alternative 5' exons was initially suggested by sequence comparisons between full length cDNAs cloned from mouse (accession #L00919) and human sources (accession #M61733). These cDNAs exhibited high nucleotide sequence homology from exon 2 through the 3' UTR, as expected for orthologous genes in human and mouse. However, the sequences upstream of exon 2 were quite distinct in these clones, consistent with the hypothesis that these sequences represent alternative first exons in the protein 4.1R gene. In order to explore this hypothesis further, we analyzed the 5' end sequences of all available 4.1R cDNAs in the various genetic databases. The strategy was to identify all clones containing bona fide 4.1R exon 2 sequences, and then classify them according to the nature of the sequences upstream of exon 2. Only those strictly adhering to the following criteria were considered to represent bona fide 4.1R 5' sequences. (1) At the mRNA level, a candidate alternative first exon must be joined, at proper exon junctions, to known downstream coding exon(s) of the 4.1R gene. (2) At the genome level, such sequences must be located in the appropriate region of human chromosome 1 upstream of the coding exons. This approach was necessary to distinguish authentic new 4.1R transcripts from potential cloning or database artifacts.

Among the collection of 4.1R cDNAs in the databases, three distinct transcript classes possessed unique 5' ends and satisfied the above criteria. The sequences comprising these 5' ends were therefore designated as exons 1A, 1B, and 1C. The sequences of the corresponding exons in mouse vs human were derived from the database clones as well as from additional 5' RACE and RT/PCR experiments. As shown in Figure 1A, the orthologous human and mouse exons are 72-83% identical at the nucleotide level. To date there is no evidence that these exons are translated; however, differential utilization of these exons does affect protein structure via coupling to downstream alternative splicing (see below).

4.1R gene organization in the 5' region, including the location of exons 1A, 1B, 1C, and 2, is shown in Figure 1B. This map, derived by analysis of BAC clones from the high throughput genomic sequence (htgs) database, revealed that the three first exons are separated by 6-20kb from one another and are located at least 70 kb upstream of the coding region. Physical linkage of these exons to the coding region was confirmed initially by the finding of two independent BAC clones containing both the exon 1 region and exon 2. More recently, these BAC sequences have been completed so that the linkage and precise distance among these various 5' exons is now known. Notably, exons 1A and 1C are located

in extensive CpG islands characteristic of many transcriptional promoter regions. Specifically, exon 1A resides in a region of 1.5kb that is 68% C+G and has a CG/GC ratio of 0.73; exon 1C is located in a 1.2 kb region that is ~69% C+G and has a CG/GC ratio of 0.80 (data not shown). Exon 1B, in contrast, is located in a region with lower C+G content (55% over 0.5kb) and exhibits a significantly lower CG/GC ratio (0.18). Nevertheless, transcription assays shown below demonstrate that exon 1B as well as 1A and 1C all possess active promoter regions.

Figure 1C summarizes the origins of the three different classes of 4.1R transcripts. According to this model, transcription can initiate at three distinct sites representing independent promoters for alternative 5' exons 1A, 1B, and 1C, each of which can be spliced directly to exon 2. It is also important to note that exon 1A splices differently to exon 2 than do exons 1B and 1C. This latter point is explored in more detail below.

Evidence for transcriptional promoter activity associated with each alternative 5' exon. The hypothesis that the widely separated exons 1A, 1B, and 1C represent bona fide alternative first exons in the 4.1R gene predicts that each must be closely associated with a distinct transcriptional promoter activity. The finding that exons 1A and 1C are located in typical CpG islands is consistent with this model. Moreover, a computational approach employing the PromoterInspector algorithm [32] also predicted transcriptional promoters corresponding almost precisely with exons 1A and 1C (results not shown). However, the most direct experimental approach was to test for promoter activity using standard luciferase reporter vectors in transfected mammalian cells. For this purpose, three 1.5kb regions of genomic DNA upstream of exons 1A, 1B, and 1C respectively, were cloned into pGL2basic upstream of the luciferase gene (Figure 2). The three candidate promoter regions exhibited strong transcriptional promoter activity relative to the promoterless luciferase reporter when tested in HEK293T cells (Figure 2). All exhibited greater than 100-fold increase in luciferase expression compared with the promoterless control, and approximately 40-80% of the activity observed with an SV40 positive control promoter. These results demonstrate robust promoter activity associated with all three exons, and strongly support the hypothesis that exons 1A, 1B, and 1C are authentic first exons in the 4.1R gene.

Evidence for coupling between alternative 5' exon choice and alternative splicing at exon 2. Close examination of 4.1R cDNAs revealed a strong correlation between the identity of the 5' exon, and the splice acceptor choice at exon 2. All reported 4.1R cDNAs that initiate at exon 1A exhibit splicing to

exon 2 at the distal (downstream) acceptor site, thus excluding the 2' region and AUG1. This conclusion was based on analysis of 20 cDNAs encompassing 7 cDNA sequences found in the databases, and 13 additional clones derived by 5'RACE from mouse spleen (data not shown). 1A transcripts would therefore be expected to encode exclusively 80kD isoforms of 4.1R protein. In contrast, all 1B transcripts identified to date, including a set of 96 products cloned from muscle by RT/PCR [11], are spliced to the proximal (upstream) acceptor site in exon 2. These transcripts include the 2' region and encode 135kD isoforms of protein 4.1R initiated at AUG1. Three cDNAs representing transcript 1C were also spliced to the proximal site and should encode 135kD protein 4.1R.

Although this correlation between mutually exclusive 5' exon use and splicing at exon 2 was intriguing, it was based on a relatively limited sample. A more systematic examination of 4.1R expression in a number of different cell types was therefore undertaken. For this approach we used RT/PCR techniques to amplify the exon 1A/2 and 1B/2 junctions, in order to distinguish between proximal and distal exon 2 splicing. Analysis of 1A transcripts was performed according to the scheme in Figure 3, which shows the predicted amplification products of 346nt or 329nt depending on whether proximal or distal splicing has occurred (upper panel). Importantly, amplification of RNA from 15 different human tissues yielded virtually identical results: in all cases only a single major PCR product was obtained, of a size consistent with splicing of exon 1A to the distal acceptor in exon 2. The identity of this product was confirmed by DNA sequence analysis. Control experiments showed that distal and proximal amplification products were resolvable under the experimental conditions employed (Figure 3A, last lane). The major conclusion from this experiment is that 1A-type transcripts in all tissues examined are spliced to the distal 3' splice site in a pattern that excludes exon 2', and thus the resulting mature mRNA will encode the 80kD forms of 4.1R protein.

Similar analysis was performed on 1B transcripts (Figure 3B). In this case, amplification of RNA from numerous tissues yielded only a single product corresponding in size and nucleotide sequence to splicing at the proximal site. These results indicate that 4.1R exon 1B transcripts splice predominantly, if not exclusively, so as to include exon 2'. Translation of 1B mature mRNA transcripts will therefore generate selectively the 135kD isoforms of 4.1R protein. This result confirms and extends the previous studies of 4.1R mRNA structure showing that mouse skeletal muscle cDNAs, with 5' ends corresponding to this newly defined exon 1B sequence, always include exon 2' [11].

Differential expression of alternative 5' exons in erythroid progenitor cells. Taken together, the data above support the idea that protein 4.1R transcripts can be expressed from one of several mutually exclusive alternative 5' exons. To explore the relative expression levels of each transcript class during late erythropoiesis, we isolated RNA from differentiating erythroblasts obtained from the spleens of mice infected with the anemia inducing strain of Friend virus [31]. These cells differentiate to the point of enucleation over the course of 48 hrs when cultured in the presence of erythropoietin, and have been shown to undergo expected changes in erythroid gene expression including induction of globin synthesis [31], upregulation of total protein 4.1R synthesis [33], and switching in protein 4.1R pre-mRNA splicing from exon 16 skipping to exon 16 inclusion [34].

Approximately equal amounts of total RNA from cells at 0, 16, 30, and 44 hrs of culture, normalized with respect to quantity of ribosomal RNA, was electrophoresed in agarose gels and subjected to Northern blot analysis with probes specific for protein 4.1R RNA (Figure 4). Hybridization with a cDNA probe to the shared coding region, that should recognize all classes of 4.1R transcript, showed that total 4.1R mRNA levels progressively increased over the course of differentiation (Figure 4A). This result is consistent with previous observations of 4.1R protein accumulation at late stages of erythropoiesis [33,35]. To determine the relative utilization of exons 1A, 1B, and 1C, Northern blotting was repeated with probes specific for each alternative 5' exon. The 1A probe hybridized to a transcript identical in size and temporal pattern of expression to that detected with the full length 4.1R probe; expression was relatively low at 0 hrs and progressively increased with differentiation time (Figure 4B). Given the finding above that 1A transcripts preferentially encode 80kD isoforms of 4.1R protein, this result is also consistent with previous observations that 80kD protein 4.1R is the predominant isoform in late erythroid cells. In contrast to the results with the 1A probe, hybridization of identical RNA blots with a probe specific for exon 1B did not detect any 4.1R mRNA corresponding to this isoform. Similarly, Northern blot analysis of mouse fetal liver total RNA revealed little or no 1B and 1C transcripts under conditions that permitted detection of abundant 1A-type 4.1R mRNA (results not shown). These results are in agreement with the low abundance of 135kD isoforms in late erythroid cells .

Differential expression of alternative 5' exons in nonerythroid cells. The RT/PCR data in Figure 3 show that exons 1A and 1B are widely expressed among human tissues, consistent with previous reports that many nonerythroid tissue express a mixture of low and high molecular weight isoforms of 4.1R protein. However, these PCR results do not give quantitative information regarding the relative

abundance of these transcripts in different cell types. The abundance of each class of transcripts was therefore examined by RNA blot analysis using short exon-specific hybridization probes. As shown in Figure 5, exons 1A and 1C were expressed in many tissues. Exon 1B was generally expressed at a low level in many tissues, but appeared relatively more abundant in cells known to exhibit features of an early erythroid phenotype (K562 cells, coordinate C10) or in tissues with significant component of early erythroid progenitors cells (fetal liver, bone marrow; coordinates D11 and G7, respectively).

**Discussion**

The results presented in this study indicate that the balance in synthesis of 135kD vs 80kD isoforms of 4.1R protein, which vary in N-terminal domain structure, is regulated by a novel mechanism involving coordination between transcriptional and alternative splicing events. Central to this model is the existence of mutually exclusive alternative 5' exons that map upstream of, and can differentially splice to, alternative splice acceptor sites flanking exon 2'. Transcripts containing exon 1A splice to the distal 3' splice site thus excluding exon 2'/AUG1 from the mature mRNA (Figure 6A). This class of 4.1R mRNA initiates translation at the downstream AUG2 and encodes 80kD isoforms of 4.1R protein (Figure 6B). In contrast, transcripts possessing the 5' exon 1B preferentially splice to the proximal 3' splice site and include AUG1, thereby generating mRNAs that encode 135kD isoforms bearing extended N-terminal domains.

According to this model, cells of both erythroid and nonerythroid origin may regulate differentially the rate of gene expression at the alternative 5' exons, as a primary mechanism for controlling their content of 4.1R protein that either includes (AUG1 isoforms) or excludes (AUG2 isoforms) the N-terminal headpiece domain. The demonstration that exons 1A, 1B, and 1C each possesses an independent transcriptional promoter supports the model that these alternative 5' sequences represent authentic first exons in the 4.1R gene, although direct mapping of 5' cap sites will be required to confirm this hypothesis. Cells with high expression of 1A transcripts would be predicted to possess a higher content of 80kD 4.1R protein (or other isoforms initiated at AUG2); conversely, cells with high levels of 1B or 1C transcripts should express more 135kD protein. In accordance with these predictions, the observed patterns of 4.1R RNA expression during erythropoiesis are consistent with previously reported qualitative and quantitative changes in the cellular content of 4.1R protein [33,35]. Early erythroid progenitors express low levels of total 4.1R protein and contain detectable amounts of 135kD isoforms

[26]. As expected, both exon 1A and exon 1B transcripts are detected in tissue sources enriched in early erythroid cells (bone marrow, fetal liver, K562 cells, Figure 5). In contrast, late erythroblasts express much higher levels of 4.1R protein that is almost exclusively of the 80kD class. Northern blot analysis of differentiating mouse erythroid cells reveals a corresponding dramatic upregulation in the expression of 1A transcripts in the absence of detectable 1B transcripts (Figure 4).

For nonerythroid cells, our results suggest that a complete characterization of 4.1R content in a given cell type will require probing at the RNA level for the presence of alternative 5' sequences, and at the protein level with antibodies that can detect both size classes of protein. Experiments performed prior to this recognition that 5' exon choice strongly influences 4.1R protein synthesis, probably underestimated the complexity of 4.1R expression in a given cell. In one nonerythroid tissue where this issue has been carefully analyzed, the kidney, there is a good correlation between the expression of exon 1A-transcripts and the synthesis of 4.1 protein initiated at AUG2; these cells express neither exon 1B nor AUG1 isoforms of 4.1R protein (P. Gascard, personal communication). The finding that most nonerythroid cells express both exon 1A- and 1B-transcripts of 4.1R is consistent with several reports demonstrating that such cells (including T lymphocytes [36], mammary epithelial cells [37], MDCK cells [38], and others [39]) do express both size classes of 4.1R protein.

In addition to the multiple promoters described here for the 4.1R gene, extensive precedence exists for multiple promoters/ multiple first exons in a number of other erythroid genes. Examples include the genes for heme biosynthetic enzymes [40-44], erythroid skeletal proteins [45-47], cell surface receptors [48], and transcription factors [49-51]. In many of these genes, the alternative first exons are exclusively noncoding; transcriptional regulation therefore controls spatial and temporal patterns of expression without altering protein structure. In a few cases, including the genes for erythroid ankyrin and band 3, the alternative promoters/ first exons are internally located, downstream of one or more coding exons, so that their respective transcripts will encode proteins with different N-termini [45-47]. The situation for 4.1R is somewhat unique: although the alternative 5' exons in the protein 4.1R gene are noncoding exons, their expression nevertheless dictates N-terminal protein structure via the coupling to downstream splicing.

Recent studies have evoked considerable interest in the coupling between transcription and several aspects of pre-mRNA processing including splicing and polyadenylation [52-55]. The mechanism by which transcriptional events may be functionally coupled with downstream alternative splicing in the

4.1R gene is not known. Pre-mRNAs initiated at alternative promoters differ in several respects and it is not yet clear which of these is most important in regulating the downstream splicing event. One critical factor may be the primary nucleotide sequence at the 5' end of each 4.1R transcript class. These unique RNA sequences might contain binding sites for splicing regulatory proteins, or might adopt specific secondary structures, either of which could affect the accessibility of downstream acceptor sites in exon 2. Another intriguing possibility is that distinct transcriptional complexes formed at the alternative 5' exons could directly affect downstream splicing events, perhaps by assembling and delivering specific splicing factors to the regulated splice sites in a co-transcriptional manner. Indeed, it has been shown that using model gene constructs that the efficiency of alternative splicing can be modulated by changes in the transcriptional promoter, even in the absence of any alteration in the sequence of the transcribed pre-mRNA itself [56,57]. Although few if any examples of coupling between transcription and alternative splicing have been reported previously in natural genes, it is likely that additional cases will be discovered as the genetic databases continue to grow. Our preliminary studies, in fact, indicate that a similar coupling of transcription and alternative splicing occurs in the protein 4.1B gene (data not shown).

Regardless of the precise mechanism, the regulated expression of 135 vs 80kD protein 4.1R isoforms is yet another example of complex expression in the 4.1R gene. An extensive array of tissue specific alternative splicing events has been characterized in the gene, including the developmentally regulated splicing of exon 16 during erythroid differentiation, the regulated expression of exon 17B in epithelial cells [37], and the preferential expression of exon 17A in muscle cells [23,37]. It seems reasonable to assume that precisely regulated expression of selected protein isoforms is indicative of an important function for these isoforms. This has already been shown in the case of exon 16, the inclusion of which results in synthesis of isoforms containing an intact spectrin-actin binding domain. While the function of the N-terminal extension in 4.1R protein isoforms is not well understood, it has been shown to influence subcellular localization [58], and may influence binding interactions of neighboring domains of the protein.

## References


1. Tchernia G, Mohandas N, Shohet SB. Deficiency of skeletal membrane protein band 4.1 in homozygous hereditary elliptocytosis. Implications for erythrocyte membrane stability. J Clin Invest. 1981;68:454-460.

2. Takakuwa Y, Tchernia G, Rossi M, Benabadji M, Mohandas N. Restoration of normal membrane stability to unstable protein 4.1-deficient membranes by incorporation of purified protein 4.1. J. Clin. Invest. 1986;78:80-85

3. Shi ZT, Afzal V, Coller B, Patel D, Chasis JA, Parra M, Lee G, Paszty C, Stevens M, Walensky L, Peters LL, Mohandas N, Rubin E, Conboy JG. Protein 4.1R-deficient mice are viable but have erythroid membrane skeleton abnormalities. J Clin Invest. 1999;103:331-340

4. Han BG, Nunomura W, Takakuwa Y, Mohandas N, Jap BK. Protein 4.1R core domain structure and insights into regulation of cytoskeletal organization. Nat Struct Biol. 2000;7:871-875.

5. Correas I. Characterization of isoforms of protein 4.1 present in the nucleus. Biochem J. 1991;279:581-585

6. De Carcer G, Lallena MJ, Correas I. Protein 4.1 is a component of the nuclear matrix of mammalian cells. Biochem J. 1995;312:871-877

7. Krauss SW, Larabell CA, Lockett S, Gascard P, Penman S, Mohandas N, Chasis JA. Structural protein 4.1 in the nucleus of human cells: dynamic rearrangements during cell division. J. Cell Biol. 1997;137:275-289

8. Lallena MJ, Correas I. Transcription-dependent redistribution of nuclear protein 4.1 to SC35-enriched nuclear domains. J Cell Sci. 1997;110:239-247

9. Krauss SW, Chasis JA, Rogers C, Mohandas N, Krockmalnic G, Penman S. Structural protein 4.1 is located in mammalian centrosomes. Proc. Natl. Acad. Sci. USA. 1997;94:7297-7302

10. Mattagajasingh SN, Huang SC, Hartenstein JS, Benz EJ, Jr. Characterization of the interaction between protein 4.1R and ZO-2. A possible link between the tight junction and the actin cytoskeleton. J Biol Chem. 2000;275:30573-30585

11. Kontrogianni-Konstantopoulos A, Huang SC, Benz EJ, Jr. A nonerythroid isoform of protein 4.1R interacts with components of the contractile apparatus in skeletal myofibers. Mol Biol Cell. 2000;11:3805-3817

12. Walensky LD, Shi ZT, Blackshaw S, DeVries AC, Demas GE, Gascard P, Nelson RJ, Conboy JG, Rubin EM, Snyder SH, Mohandas N. Neurobehavioral deficits in mice lacking the erythrocyte membrane cytoskeletal protein 4.1. Curr Biol. 1998;8:1269-1272

13. Nunomura W, Takakuwa Y, Tokimitsu R, Krauss SW, Kawashima M, Mohandas N. Regulation of CD44-protein 4.1 interaction by Ca2+ and calmodulin. Implications for modulation of CD44-ankyrin interaction. J Biol Chem. 1997;272:30322-30328

14. Hemming NJ, Anstee DJ, Staricoff MA, Tanner MJ, Mohandas N. Identification of the membrane attachment sites for protein 4.1 in the human erythrocyte. J. Biol. Chem. 1995;270:5360-5366

15. Marfatia SM, Leu RA, Branton D, Chishti AH. Identification of the protein 4.1 binding interface on glycophorin C and p55, a homologue of the Drosophila discs-large tumor suppressor protein. J Biol Chem. 1995;270:715-719.

16. Cohen AR, Wood DF, Marfatia SM, Walther Z, Chishti AH, Anderson JM. Human CASK/LIN-2 Binds Syndecan-2 and Protein 4.1 and Localizes to the Basolateral Membrane of Epithelial Cells. J Cell Biol. 1998;142:129-138

17. Lue RA, Marfatia SM, Branton D, Chishti AH. Cloning and characterization of hdlg: The human homologue of the *Drosophila* discs large tumor suppressor binds to protein 4.1. Proc. Natl. Acad. Sci. USA. 1994;91:9818-9822

18. Hou CL, Tang C, Roffler SR, Tang TK. Protein 4.1R binding to eIF3-p44 suggests an interaction between the cytoskeletal network and the translation apparatus. Blood. 2000;96:747-753

19. Hung LY, Tang CJ, Tang TK. Protein 4.1 R-135 interacts with a novel centrosomal protein (CPAP) which is associated with the gamma-tubulin complex. Mol Cell Biol. 2000;20:7813-7825

20. Lallena MJ, Martinez C, Valcarcel J, Correas I. Functional association of nuclear protein 4.1 with pre-mRNA splicing factors. J Cell Sci. 1998;111:1963-1971

21. Tang CJ, Tang TK. The 30-kD domain of protein 4.1 mediates its binding to the carboxyl terminus of pICln, a protein involved in cellular volume regulation. Blood. 1998;92:1442-1447

22. Huang JP, Tang CJ, Kou GH, Marchesi VT, Benz EJ, Jr., Tang TK. Genomic structure of the locus encoding protein 4.1. Structural basis for complex combinational patterns of tissue-specific alternative RNA splicing. J Biol Chem. 1993;268:3758-3766

23. Baklouti F, Huang SC, Vulliamy TJ, Delaunay J, Benz EJ, Jr. Organization of the human protein 4.1 genomic locus: new insights into the tissue-specific alternative splicing of the pre-mRNA. Genomics. 1997;39:289-302

24. Conboy J. The role of alternative pre-mRNA splicing in regulating the structure and function of skeletal protein 4.1. Proc Soc Exp Biol Med. 1999;220:73-78

25. Gascard P, Lee G, Coulombel L, Auffray I, Lum M, Parra M, Conboy JG, Mohandas N, Chasis JA. Characterization of multiple isoforms of protein 4.1R expressed during erythroid terminal differentiation. Blood. 1998;92:4404-4414

26. Chasis JA, Coulombel L, Conboy J, McGee S, Andrews K, Kan YW, Mohandas N. Differentiation-associated switches in protein 4.1 expression. Synthesis of multiple structural isoforms during normal human erythropoiesis. J Clin Invest. 1993;91:329-338

27. Baklouti F, Huang SC, Tang TK, Delaunay J, Marchesi VT, Benz EJ, Jr. Asynchronous regulation of splicing events within protein 4.1 pre-mRNA during erythroid differentiation. Blood. 1996;87:3934-3941

28. Gee SL, Aoyagi K, Lersch R, Hou V, Wu M, Conboy JG. Alternative splicing of protein 4.1R exon 16: ordered excision of flanking introns ensures proper splice site choice. Blood. 2000;95:692-699

29. Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene. 1989;77:61-68

30. Sawyer ST, Koury MJ, Bondurant MC. Large-scale procurement of erythropoietin-responsive erythroid cells: assay for biological activity of erythropoietin. Methods Enzymol. 1987;147:340-352

31. Koury MJ, Sawyer ST, Bondurant MC. Splenic erythroblasts in anemia-inducing Friend disease: a source of cells for studies of erythropoietin-mediated differentiation. J Cell Physiol. 1984;121:526-532

32. Scherf M, Klingenhoff A, Werner T. Highly specific localization of promoter regions in large genomic sequences by PromoterInspector: a novel context analysis approach. J Mol Biol. 2000;297:599-606.

33. Koury MJ, Bondurant MC, Mueller TJ. The role of erythropoietin in the production of principal erythrocyte proteins other than hemoglobin during terminal erythroid differentiation. J Cell Physiol. 1986;126:259-265

34. Hou VC, Gee SL, Wu M, Turck CW, Koury MJ, Krainer AR, Mayeda A, Conboy JC. Two hnRNP A/B-dependent exonic splicing silencer elements are required to repress 4.1R exon 16 splicing. Blood. 2001;98, Suppl 1:436a

35. Hanspal M, Hanspal JS, Kalraiya R, Liu SC, Sahr KE, Howard D, Palek J. Asynchronous synthesis of membrane skeletal proteins during terminal maturation of murine erythroblasts. Blood. 1992;80:530-539

36. Conboy JG, Chan JY, Chasis JA, Kan YW, Mohandas N. Tissue- and development-specific alternative RNA splicing regulates expression of multiple isoforms of erythroid membrane protein 4.1. J Biol Chem. 1991;266:8273-8280

37. Schischmanoff PO, Yaswen P, Parra MK, Lee G, Chasis JA, Mohandas N, Conboy JG. Cell shape-dependent regulation of protein 4.1 alternative pre-mRNA splicing in mammary epithelial cells. J Biol Chem. 1997;272:10254-10259

38. Mattagajasingh SN, Huang SC, Hartenstein JS, Snyder M, Marchesi VT, Benz EJ. A nonerythroid isoform of protein 4.1R interacts with the nuclear mitotic apparatus (NuMA) protein. J Cell Biol. 1999;145:29-43

39. Anderson RA, Correas I, Mazzucco C, Castle JD, Marchesi VT. Tissue-specific analogues of erythrocyte protein 4.1 retain functional domains. J Cell Biochem. 1988;37:269-284

40. Bishop TR, Miller MW, Beall J, Zon LI, Dierks P. Genetic regulation of delta-aminolevulinate dehydratase during erythropoiesis. Nucleic Acids Res. 1996;24:2511-2518

41. Kaya AH, Plewinska M, Wong DM, Desnick RJ, Wetmur JG. Human delta-aminolevulinate dehydratase (ALAD) gene: structure and alternative splicing of the erythroid and housekeeping mRNAs. Genomics. 1994;19:242-248

42. Grandchamp B, De VH, Beaumont C, Chretien S, Walter O, Nordmann Y. Tissue-specific expression of porphobilinogen deaminase. Two isoenzymes from a single gene. Eur J Biochem. 1987;162:105-110

43. Romeo PH, Mignotte V, Raich N, Dubart A, Beaupain D, Romana M, Chabret C, Wall L, deBoer E, Grosveld F. Regulated expression of the erythroid-specific-promoter of the human PBG-D gene during erythroid differentiation. Prog Clin Biol Res. 1989:359-366

44. Aizencang GI, Bishop DF, Forrest D, Astrin KH, Desnick RJ. Uroporphyrinogen III synthase. An alternative promoter controls erythroid-specific expression in the murine gene. J Biol Chem. 2000;275:2295-2304

45. Sahr KE, Daniels BP, Hanspal M. Identification of the proximal erythroid promoter region of the mouse anion exchanger gene. Blood. 1996;88:4500-4509

46. Gallagher PG, Romana M, Tse WT, Lux SE, Forget BG. The human ankyrin-1 gene is selectively transcribed in erythroid cell lines despite the presence of a housekeeping-like promoter. Blood. 2000;96:1136-1143.

47. Birkenmeier CS, Sharp JJ, Gifford EJ, Deveau SA, Barker JE. An alternative first exon in the distal end of the erythroid ankyrin gene leads to production of a small isoform containing an NH2-terminal membrane anchor. Genomics. 1998;50:79-88

48. Chin K, Yu X, Beleslin-Cokic B, Liu C, Shen K, Mohrenweiser HW, Noguchi CT. Production and processing of erythropoietin receptor transcripts in brain. Brain Res Mol Brain Res. 2000;81:29-42

49. Moroni E, Cairns L, Ottolenghi S, Giglioni B, Ashihara E, Migliaccio G, Migliaccio AR. Expression in hematopoietic cells of GATA-1 transcripts from the alternative "testis" promoter during development and cell differentiation. Biochem Biophys Res Commun. 1997;231:299-304

50. Vannucchi AM, Linari S, Lin CS, Koury MJ, Bondurant MC, Migliaccio AR. Increased expression of the distal, but not of the proximal, Gata1 transcripts during differentiation of primary erythroid cells. J Cell Physiol. 1999;180:390-401

51. Toki T, Arai K, Terui K, Komatsu N, Yokoyama M, Katsuoka F, Yamamoto M, Ito E. Functional characterization of the two alternative promoters of human p45 NF-E2 gene. Exp Hematol. 2000;28:1113-1119

52. Bentley D. Coupling RNA polymerase II transcription with pre-mRNA processing. Curr Opin Cell Biol. 1999;11:347-351.

53. Hirose Y, Manley JL. RNA polymerase II and the integration of nuclear events. Genes Dev. 2000;14:1415-1429

54. Proudfoot N. Connecting transcription to messenger RNA processing. Trends Biochem Sci. 2000;25:290-293.

55. Maniatis T, Reed R. An extensive network of coupling among gene expression machines. Nature. 2002;416:499-506.

56. Cramer P, Pesce CG, Baralle FE, Kornblihtt AR. Functional association between promoter structure and transcript alternative splicing. Proc Natl Acad Sci U S A. 1997;94:11456-11460

57. Cramer P, Caceres JF, Cazalla D, Kadener S, Muro AF, Baralle FE, Kornblihtt AR. Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. Mol Cell. 1999;4:251-258

58. Luque CM, Lallena MJ, Perez-Ferreiro CM, de Isidro Y, De Carcer G, Alonso MA, Correas I. The N-terminal 209-aa domain of high molecular-weight 4.1R isoforms abrogates 4.1R targeting to the nucleus. Proc Natl Acad Sci U S A. 1999;96:14925-14930

**Figure legends**

Figure 1. **Alternative 5' exons of the 4.1R gene**. **A.** Sequence comparison of human vs mouse alternative 4.1R 5' exons. The orthologous exons had the following sequence identities: 1A, 72% over 142nt; 1B, 79% over a region of 107nt; 1C, 83% over 159nt. Origins of sequences: for exon 1A, mouse sequence was from AA014918, and human from M61733 and BG943340. For exon 1B, mouse sequence was from L00919, while the human sequence was deduced by comparing the mouse 1B sequence to human genomic sequence downstream of exon 1A. That the predicted human 1B can splice to exon 2 was confirmed by RT/PCR analysis (Figure 3). For exon 1C, human sequence was from AL041809 while mouse sequence was derived from a 5' RACE product from mouse spleen (results not shown). **B.** Structure of the 5' region of the 4.1R gene. The genomic map was constructed from overlapping BAC human genomic clones derived from high throughput genomic sequences. Distances between exons (in kilobases) are indicated, and the location of CpG islands is shown. **C**. Model showing independent splicing of alternative 5' exons to downstream exon 2, to generate three 4.1R transcript classes with different 5' end sequences.

Figure 2. **Transcriptional promoter activity associated with exons 1A, 1B, and 1C**. Shown is firefly luciferase activity assayed 48hrs after transfection of HEK293T cells, relative to the activity of cells transfected in parallel with the promoterless pGL2 basic control vector. Assays were performed in triplicate and transfection efficiency was normalized with respect to co-transfected Renilla luciferase activity in the same cultures. As a positive control, luciferase activity expressed from a standard SV40 promoter is shown.

Figure 3. **Coupling between alternative 5' exons and alternative splicing at exon 2**. Upper panel: 1A transcripts were amplified from the indicated human tissues using primers in exons 1A and 2. Control shows migration of PCR products that include exon 2' (346nt; proximal splicing) or exclude exon 2' (329nt; distal splicing). Lower panel: 1B transcripts were amplified from the

indicated tissues using primers in exons 1B and 2. Control shows migration of PCR products that include exon 2' (295nt; proximal splicing) or exclude exon 2' (278nt; distal splicing).

Figure 4. **Expression of alternative 5' exons in erythroid progenitor cells**. Shown is a Northern blot analysis of total RNA from FVA-induced erythroblasts following culture for the times indicated (in hours). Identity of hybridization probe is shown at left of each blot. A. Full length cDNA probe. B. Exon 1A probe. Both probes detected a single 4.1R mRNA band that increased during the course of differentiation. No hybridization was detected with probes for exon 1B and 1C under identical conditions (not shown). C. Ethidium bromide staining of ribosomal RNA. Equal (0, 16, and 30hr) or slightly reduced (44hr) loads of RNA were applied for each time point, supporting the idea that a true increase in 4.1R RNA expression has occurred in these cells.

Figure 5. **Tissue-specific expression of alternative 5' exons**. A master RNA blot (Clontech) containing poly A+ RNA from many human tissues and cell lines, was blotted with probes for alternative exon 1A (panel A), exon 1B (panel B), or exon 1C (panel C). 1A and 1C transcripts were particularly abundant in tissues with a high content of muscle cells (heart, A4; skeletal muscle, B7; stomach, B5) as well as in pancreas (B9) and liver (A9). In contrast, exon 1B transcripts were most abundant in tissues enriched for early erythroid progenitors (fetal liver, D11; bone marrow, G7) and the erythroleukemia cell line K562 (C10). Panel D shows the tissue map for all RNA sources tested.

Figure 6. **Model of protein 4.1R 5' gene expression**. A. Upper panel shows differential splicing of 5' putative exons to alternative splice acceptor sites in exon 2. B. Lower panel shows the distinct protein isoforms resulting from translation of 1A- and 1B-type mRNAs. MBD, membrane binding domain; U2, unique region 2; SAB, spectrin-actin binding domain; CTD, C-terminal domain; ss, splice site.

Figure 1

## exon 1A

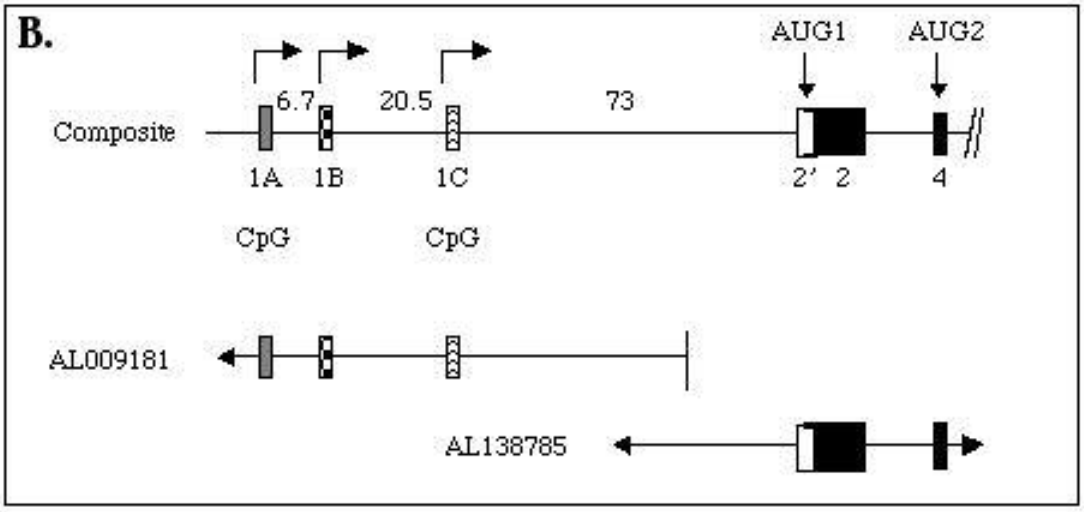
```
human  gttgccctgtcagtgcaggtggaggccccggcggggcaaagtggcaggaacctcttaaagggcgagagcggcgcg
       |||||||||||||||||||||||||||| |||||  | |||||||||||||||||||||| |||||| |||
mouse  gttgccctgtcagtgcaggtggaggcccccgcgggcaagagtggcaggaacctcttaaagggagagagccgcgga

human  gagccagaacgcgg-tcggcccggtccccgccgcacc----------------------cagcccag
       |||| || ||||||  |||||||| ||| |||    |                      ||||||||
mouse  gagcgaggacgcggcccggcccgg-cccggccaggtccctgccaggtccctgtcctgtgcagcccag
```

## exon 1B

```
human  gctgtgctgtgtgtctcactgctgtcatttgcttccatgtgccactgcggtgatggctcca-tggcctct
       ||||||||||||  ||||||| ||||| |||||||||||||||||| | |  ||||||| |  || ||| |
mouse  gctgtgctgtgt--ctcacttctgtcctttgcttccatgtgccaccgtgacgatggctgcgctgtcct-t

human  cctggactggctgcgtgaccccgacggctgcttctgg
       ||| |||||||||  |||   | ||| || |||||||
mouse  ccttgactggctgtatgatgacaacgactacttctgg
```

## exon 1C

```
human  agggcgcgcgccagggtcgctcgctcgctccctccctccgctggtgcgtttagtcagtcagccagcagcagccgg
       ||||| || ||    | ||        |||||||||||||| || |  |||| |||| ||||| ||   |||| |
mouse  agggctcgagcggccgcccgggcaggtctccctccctccgccggcggctttactcagccagcctgc---agcccg

human  cgggcccgcgagccgcagagggcccgagcctcggacgccggcgcctcctcctgccattgttcgtcgggctgcagc
       |||||||| |||||||||||||| |||||||||||| |||||| |||||||||||||||||||||||||| |||
mouse  ccagcccgggagccgcagagggccggagcctcggacgcgggcgccgcctcctgccattgttcgtcgggctgaagc

human  agtggcgggcgcaggagcccggccccggagccaccg
        | ||| ||||| ||||||| | ||  || || | |
mouse  tgcggccggcgcgggagccccggccgagaaccccgg
```


B.
AUG1
AUG2
6.7
20.5
73
Composite
1A
1B
1C
2'
2
4
CpG
CpG
AL009181
AL138785

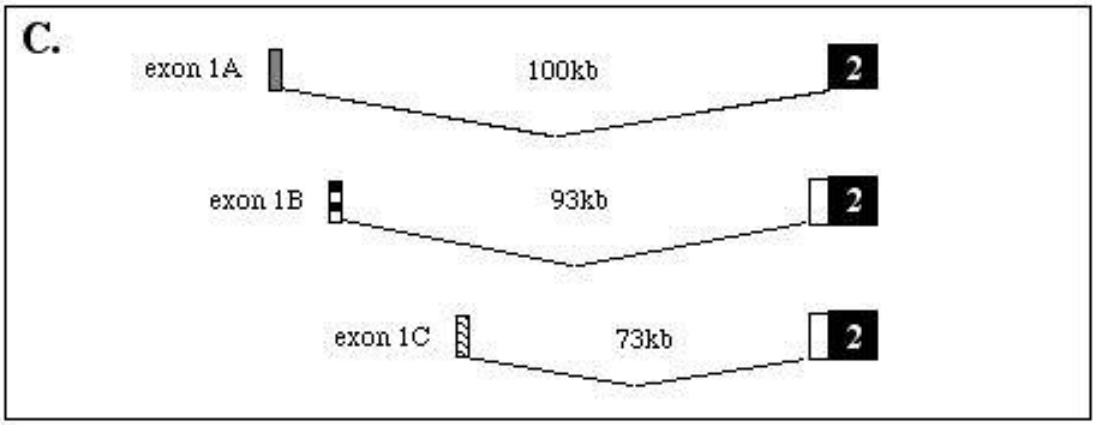

C.
exon 1A
100kb
2
exon 1B
93kb
2
exon 1C
73kb
2

Figure 1B, 1C

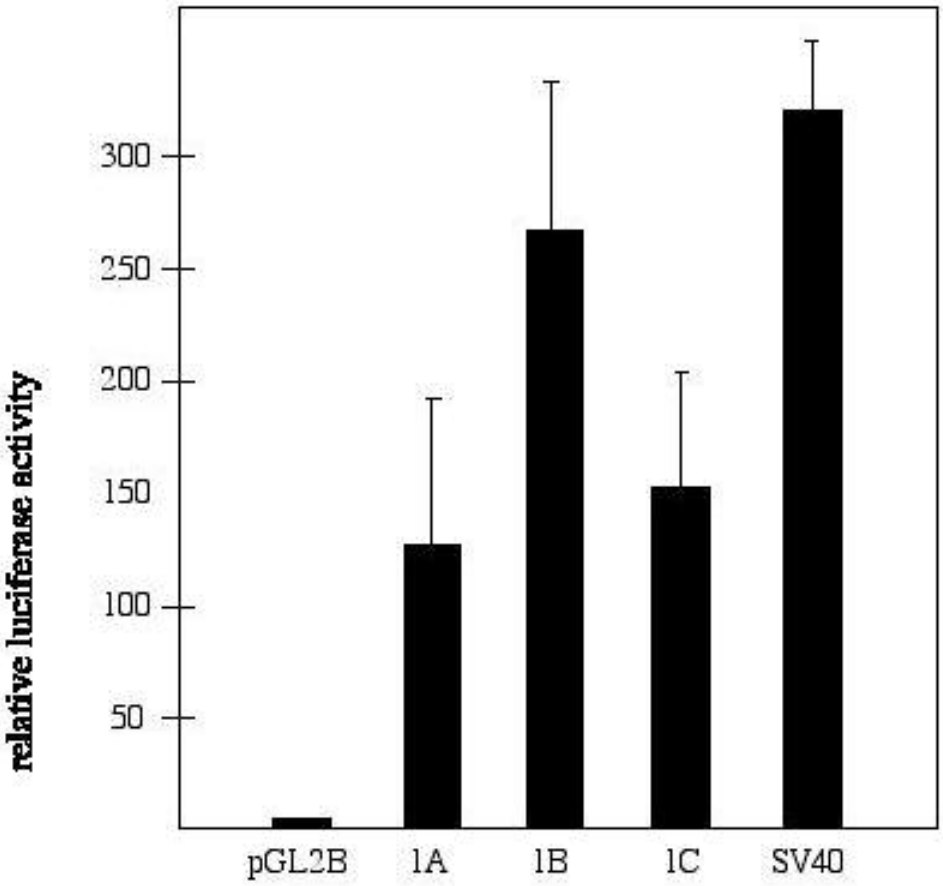

relative luciferase activity
300
250
200
150
100
50
pGL2B
1A
1B
1C
SV40

Figure 2

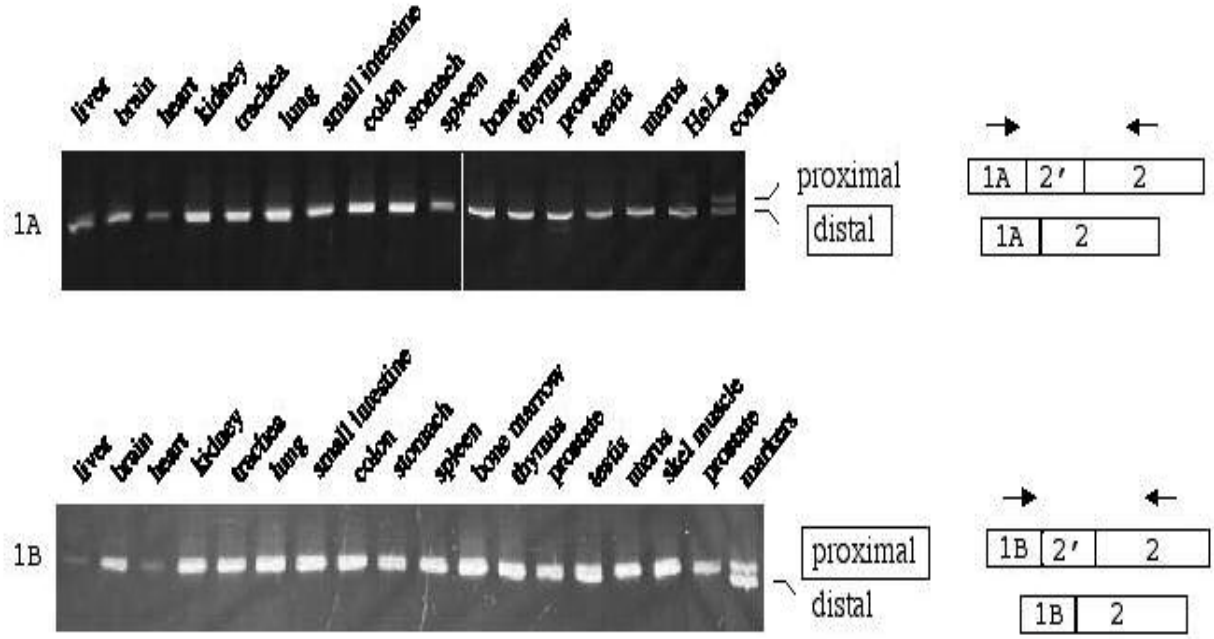

liver
brain
heart
kidney
trachea
lung
small intestine
colon
stomach
spleen
bone marrow
thymus
prostate
testis
uterus
HeLa
controls
1A
proximal
distal
1A
2′
2
1A
2
liver
brain
heart
kidney
trachea
lung
small intestine
colon
stomach
spleen
bone marrow
thymus
prostate
testis
uterus
skel muscle
prostate
markers
1B
proximal
distal
1B
2′
2
1B
2

Figure 3

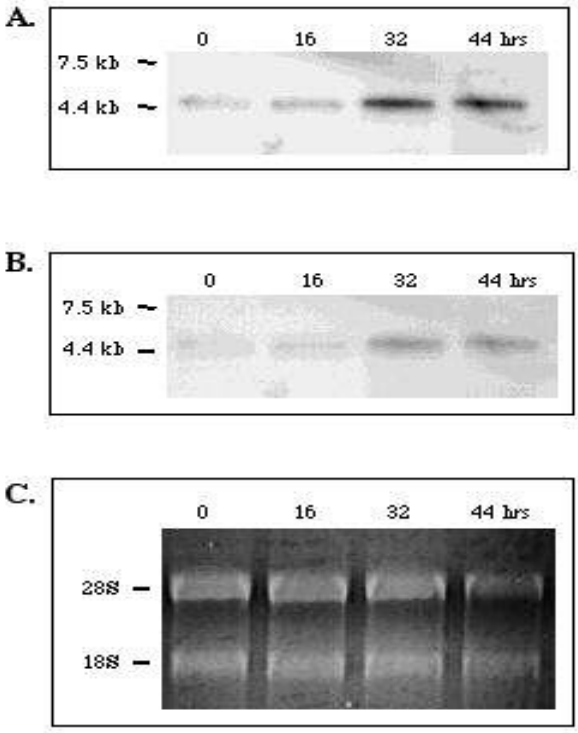

A.
0
16
32
44 hrs
7.5 kb
4.4 kb
B.
0
16
32
44 hrs
7.5 kb
4.4 kb
C.
0
16
32
44 hrs
28S
18S

Figure 4

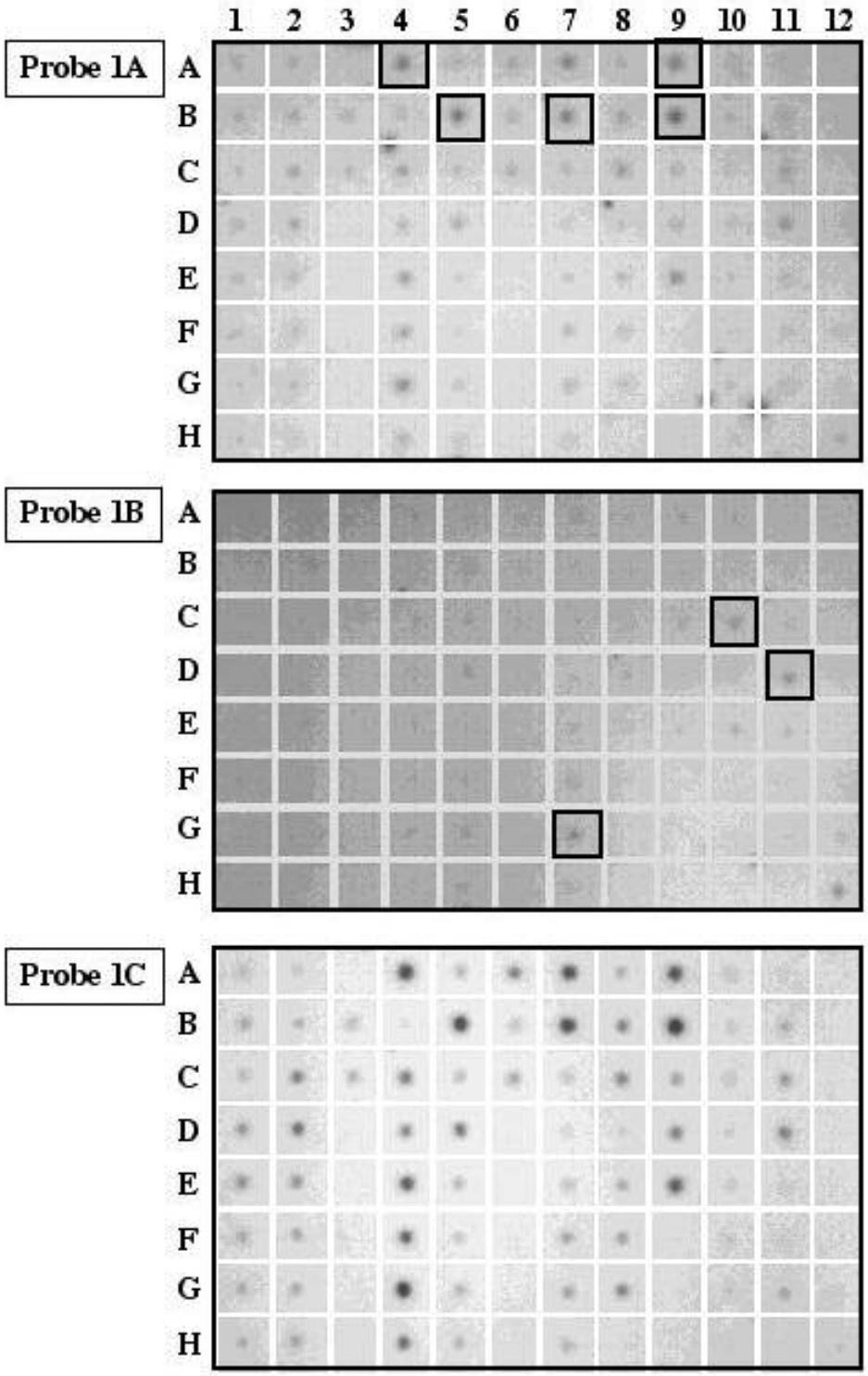

1 2 3 4 5 6 7 8 9 10 11 12
Probe 1A
A
B
C
D
E
F
G
H
Probe 1B
A
B
C
D
E
F
G
H
Probe 1C
A
B
C
D
E
F
G
H

Figure 5

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | whole brain | cerebellum left | substantia nigra | heart | esophagus | colon, transverse | kidney | lung | liver | leukemia, HL-60 | fetal brain | yeast total RNA |
| B | cerebral cortex | cerebellum right | nucleus accumbens | aorta | stomach | colon, descending | skeletal muscle | placenta | pancreas | HeLa S3 | fetal heart | yeast tRNA |
| C | frontal lobe | corpus callosum | thalamus | atrium left | duodenum | rectum | spleen | bladder | adrenal gland | leukemia K-562 | fetal kidney | Ecoli tRNA |
| D | parietal lobe | amygdala | pituitary gland | atrium right | jejunum | | thymus | uterus | thyroid gland | leukemia MOLT-4 | fetal liver | Ecoli DNA |
| E | occipital lobe | caudate nucleus | spinal cord | ventricle left | ileum | | peripheral blood leukocyte | prostate | salivary gland | Burkitt's lymphoma Raji | fetal spleen | poly r(A) |
| F | temporal lobe | hippo-campus | | ventricle right | ileocecum | | lymph node | testis | mammary gland | Burkitt's lymphoma Daudi | fetal thymus | human C.t-1 DNA |
| G | p.g. of cerebral cortex | medulla oblongata | | inter-ventricular septum | appendix | | bone marrow | ovary | | colorectal adenocarcinoma SW480 | fetal lung | human DNA 100ng |
| H | pons | putamen | | apex of the heart | colon, ascending | | trachea | | | lung carcinoma 549 | | human DNA 500ng |

Figure 5, continued

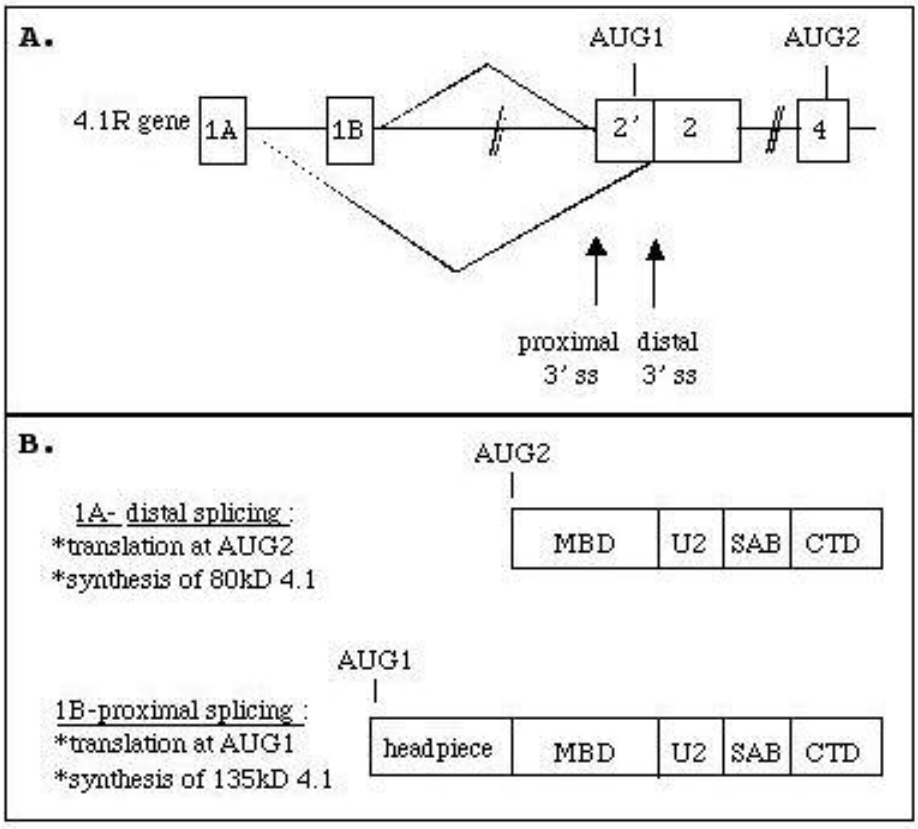

A.
AUG1
AUG2
4.1R gene
1A
1B
2'
2
4
proximal 3' ss
distal 3' ss
B.
AUG2
1A- distal splicing :
*translation at AUG2
*synthesis of 80kD 4.1
MBD
U2
SAB
CTD
AUG1
1B-proximal splicing :
*translation at AUG1
*synthesis of 135kD 4.1
headpiece
MBD
U2
SAB
CTD

Figure 6